AN ANALYSIS OF THE NAVAL PERSONNEL PAY PREDICTOR (ENLISTED MODEL)

Allan Ray Walker

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

AN ANALYSIS OF THE
NAVAL PERSONNEL PAY PREDICTOR
(ENLISTED MODEL)

by

Allan Ray Walker

September 1975

Thesis Advisor: R.W. Butterworth

Approved for public release; distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)*
An Analysis of the Naval Personnel Pay Predictor (Enlisted Model)

5. TYPE OF REPORT & PERIOD COVERED
Master's Thesis;
September 1975

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Allan Ray Walker

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93940

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Naval Postgraduate School
Monterey, California 93940

12. REPORT DATE
September 1975

13. NUMBER OF PAGES
52

14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*

15. SECURITY CLASS. *(of this report)*
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*
Enlisted Pay Predictor

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The Naval Personnel Pay Predictor (Enlisted Model) is used by the Bureau of Naval Personnel as a tool for predicting the total annual basic pay for the enlisted force as an input to the budget process. A major source of error in the model was found to be the prediction of the length of service (LOS) vector, and an attempt to improve this prediction was made. The extreme complexity of the model was found to be unnecessary,

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

(20. ABSTRACT Continued)

and a simple exponential smoothing subroutine for LOS prediction did as well or better than the original model. It was also found that a double exponential smoothing subroutine, taking into account the trends in the force structure, would almost uniformly improve the one year prediction from the model.

An Analysis of the
Naval Personnel Pay Predictor
(Enlisted Model)

by

Allan Ray Walker
Lieutenant, United States Navy
B.S., University of Louisville, 1968
M.S., University of West Florida, 1970

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1975

ABSTRACT

The Naval Personnel Pay Predictor (enlisted Model) is used by the Bureau of Naval Personnel as a tool for predicting the total annual basic pay for the enlisted force as an input to the budget process. A major source of error in the model was found to be the prediction of the length of service (LOS) vector, and an attempt to improve this prediction was made. The extreme complexity of the model was found to be unnecessary, and a simple exponential smoothing subroutine for LOS prediction did as well or better than the original model. It was also found that a double exponential smoothing subroutine, taking into account the trends in the force structure, would almost uniformly improve the one year prediction from the model.

TABLE OF CONTENTS

I. NATURE OF THE PROBLEM ---------------------------- 7

II. THE NAPPE MODEL ---------------------------------- 8

A. SMOOTHING THE LOS VECTORS OF THE INVENTORIES - 8

B. GENERATING THE PAY GRADE BY LOS MATRIX ------- 11

III. EXPERIMENTS WITH AND CHANGES TO THE NAPPE MODEL -- 13

A. MAJOR SOURCE OF ERROR ------------------------ 13

B. SIMPLE EXPONENTIAL SMOOTHING ----------------- 15

1. The SMOTHY Subroutine -------------------- 16

2. Results of SMOTHY ------------------------ 19

C. DOUBLE EXPONENTIAL SMOOTHING ----------------- 20

1. The SMOTH2 Subroutine -------------------- 21

2. Results of SMOTH2 ------------------------ 22

D. THE MATRIX GENERATION PROCEDURE -------------- 24

IV. CONCLUSION ------------------------------------- 26

APPENDIX A --- 28

APPENDIX B --- 32

COMPUTER PROGRAM -- 37

LIST OF REFERENCES -------------------------------------- 50

INITIAL DISTRIBUTION LIST ------------------------------- 51

LIST OF TABLES

I. Comparison of NAPPE with Pure Mosteller -------- 14

II. Comparison of NAPPE and SMOTHY Predictions ----- 19

III. Comparison of Absolute Percentage Errors ------- 23

I. NATURE OF THE PROBLEM

The Bureau of Naval Personnel requires many mathematical models for accurately predicting the structure of the future force. These models are used as tools to aid in planning decisions. Of special interest is the problem of costing the future force as a part of the budget submission procedure.

The Bureau of Naval Operations determines the personnel requirements for the future force and passes these to the Bureau of Personnel for implementation. These requirements are presented in the form of quarterly pay grade vectors, that is, the number of people required in pay grade E-1, E-2, E-3, ..., E-9. Since the amount of pay received is dependent on the member's length of service, the problem of predicting the total cost of the force becomes complex.

The specific problem considered was: given the future size of the force by pay grade, the past and present inventories, predict the total annual base pay of the force for future years.

II. THE NAPPE MODEL

One model currently used for this purpose is the Naval Personnel Pay Predictor (Enlisted Model), referred to as NAPPE. The model makes use of a data base consisting of three sets of quarterly inventories (pay grade by LOS Force matrices) for all years since 1957. The inventories are for United States Navy (USN), United States Naval Reserve (USNR), and Total Navy (TOTALNAV), the sum of the two.

The procedure is to, first, predict the future quarterly LOS vectors for the desired number of years into the future (up to ten). This is a vector of the total number of people with length of service 1, 2, ..., 31 years. The methodology used for this prediction is discussed later in this section and in Appendices A and B. The LOS vector is then combined with the pay grade requirements vector to get the predicted force matrix. A discussion of this procedure is also included in this section. The cost of the force is then simply the multiplication of the straight line averaged (between successive quarters) number of people in each cell of the matrix with the pay scale for that cell, which is an input to the model.

A. SMOOTHING THE LOS VECTORS OF THE INVENTORIES

The first step in the NAPPE models prediction of the LOS vector is accomplished by a subroutine referred to as SMOOTH

(refer to Appendix A for the mathematical model). Throughout the discussion of SMOOTH it should be remembered that all calculations applied to previous years data are made independently for the population (the total number of people) in each element of the LOS vector (hence referred to as LOS cell) and the transition rates from one cell to the next, computed for all three data bases. The transition rate is simply the proportion of the population in cell i of year j which move to cell i+1 in year j+1. The methodology is basic single exponential smoothing as discussed by Brown [1] and others.

The following procedure is done independently for each LOS cell. For each year of historical data, a prediction is made based on exponentially smoothing the data up to that year using values of the smoothing constant (hence referred to as alpha) of .05, .10, ..., .95. For each year, the predicted value is then compared with the actual value to determine which value of alpha would have given the best prediction. This results in the selection of an alpha for each LOS cell for each year of data. Consult Appendix A for the exact procedure and forms of the resulting error that are stored and used by the model. "Best" predictions and resulting errors are made for all years of historical data, finally resulting in a decision for the "best" alpha for predicting the future.

The output of SMOOTH consists of four sets of LOS vectors. For each year of historical data, there is a prediction based

on transition rates and a prediction based on previous year's cell populations, one pair based on the TOTALNAV data and the other pair on the sum of the predictions based on USN and USNR data. Note that due to the difference in the structure of the USN and USNR, the sum of these predictions may be different than the prediction based on their sum.

The final prediction is made by a subroutine referred to as ADJSMO (refer to Appendix B for the model). ADJSMO considers five "methods" for prediction. These include the four outputs from SMOOTH plus a weighted average of these predictions. This weighted average is formed by multiplying a weight (BWT) times the average of the two transition rate based predictions plus the complementary weight (1.0 - BWT) times the average of the two population based predictions. This calculation is made for values of BWT of .45, .50, ..., .95. For each year of data a "best" method (of the five), in the least square error sense, and a "best" weight, if a weighted average method was chosen, is selected for predicting that year. The absolute sum of the errors of the "best" prediction is also calculated for use in adjusting the final prediction. This adjustment is necessary because no transition rate if available to predict LOS cell 1.

Having selected a "best" method and a "best" weighting factor based on the last year of historical data, the model predicts the first future year values for LOS cells 2-31. At this point the model calculates the average (over all

years of data) proportion of the total population that was in LOS cell 1. This proportion is then applied to the total force required for the quarter under consideration and compared with the number which would be in cell 1 given the predicted values for cells 2-31 and the required total. Half of the difference in these two values is then allocated among cells 2-31 according to the total absolute error discussed above. The prediction for cell 1 is then the difference between the required total size of the force and the predictions made for cells 2-31.

B. GENERATING THE PAY GRADE BY LOS MATRIX

The pay grade by LOS matrix is calculated using a method for renormalizing contingency tables, as described by Mosteller [Ref. 3]. This method is an iterative procedure which takes the desired marginal totals of a matrix and a given, or base, matrix of the desired form and constructs a matrix as similar as possible to the base matrix, having the marginal totals that were desired.

Since this method was used throughout the research, a brief discussion of the procedure follows:

Let $A_{i,j}$ be the elements of the base matrix, $i = 1,2,...,9$, $j = 1,2,...,31$

Let R_j be the desired row totals

Let C_i be the desired column totals

$$(1) \quad R'_j = \sum_{i=1}^{9} A_{i,j} \qquad \text{for all } j$$

$$D_j = R_j/R'_j \qquad \text{for all } j$$

$$A'_{i,j} = D_j\ A_{i,j} \qquad \text{for all } i,j$$

$$C'_i = \sum_{j=1}^{31} A'_{i,j} \qquad \text{for all } i$$

$$D'_i = C_i/C'_i \qquad \text{for all } i$$

$$A''_{i,j} = D'_i\ A'_{i,j} \qquad \text{for all } i,j$$

$$A_{i,j} = A''_{i,j}$$

Return to step (1).

The procedure is continued until the row and column totals converge to the desired totals.

In the NAPPE model, the marginals are the given pay grade vector and the predicted LOS vector. The base matrix is calculated as the simple average of the last twelve quarterly historic inventories.

III. EXPERIMENTS WITH AND CHANGES TO THE NAPPE MODEL

Since the object of the research was to improve the predictive accuracy of the NAPPE model, the sources of error had to be determined. It appeared that there were two independent sources of error, predicting the LOS vector and the instability of the Mosteller procedure for completing the matrix. Both of these possible problem areas were studied. In this section is a discussion of the first area and changes which were made to the model to improve its predictive quality. This is followed by a discussion of a study to discover the factors which influence the Mosteller procedure.

A. MAJOR SOURCE OF ERROR

The removal of either of these above mentioned sources of error should improve the predictions of the model. An ad hoc test of this hypothesis was accomplished by using the Mosteller subroutine (called PNGPNG) with the correct LOS vector and comparing the results of the model with known values, for years with historical data available.

The NAPPE model has a validation feature which facilitates this and other kinds of comparisons. As an input to the model, the last date of historical data to be used is given. The model then only looks at data up to that date and predicts as if that were today's date. Also included in the NAPPE package (which consists of several minor models besides

NAPPE itself) is a model called NAPVAL. This model compares the NAPPE output with the actual inventories. These comparisons are discussed throughout this paper. Specifically, any number called "actual" will mean an output from NAPVAL. Also, throughout the paper, the measure of effectiveness for comparison will be the total annual cost of the force, which is an output of both models.

In order to accomplish the above mentioned objective, the SMOOTH and ADJSMC subroutines were removed from the model. In their place, the actual LOS vector was read from the inventories and the following table is the result of comparing the prediction based on this procedure and the prediction of NAPPE. The elements of the table are the actual cost (NAPVAL), the NAPPE prediction (with the model untouched), and the prediction using only the Mosteller procedure (NAPPE with SMOOTH and ADJSMO removed), labeled "Using Actual LOS".

TABLE I

COMPARISON OF NAPPE WITH PURE MOSTELLER
(costs are in millions of dollars)

Year	Actual Cost	NAPPE Prediction	% Error	Using Actual Los	% Error
1968	1,832	1,839	.363	1,832	.014
1969	2,009	2,016	.372	2,010	.061
1970	2,280	2,286	.262	2,282	.081
1971	2,264	2,258	.290	2,266	.092

These results indicated that a major source of error in the model, and hence a potential for improvement, results from the prediction of the LOS vector, as was expected. If one looks at any feature of the enlisted force (such as size or distribution), he finds that it is not stationary in time, even considering statistical fluctuations. There are obvious trends. During war years, the force becomes larger and, on the average, younger, while during peace time, the force becomes smaller and older. Since single exponential smoothing does not allow for trends, it could not be expected to handle the problem being considered.

However, before attacking this problem, there were other questions to be answered. After documenting the model, two other questions came to mind. Is the pure complexity of the model worth the computer requirements? (The following section indicates not.) Is the use of the entire data base justified? Intuitively, the answer to the second question was no. The size and structure of the force in the late 1950's is not indicative of the force in 1975. There are continual policy changes which affect enlistment, promotion, and retention.

B. SIMPLE EXPONENTIAL SMOOTHING

In order to answer these questions, the first change in the model was made. The SMOOTH and ADJSMO subroutines were removed and a subroutine, SMOTHY, replaced them. This

subroutine used simple exponential smoothing of the transition rates and only the TOTALNAV data base.

1. <u>The SMOTHY Subroutine</u>

This subroutine results in an extensive simplification of the NAPPE model as it uses only four years of historical data and a single alpha value of 0.4. This alpha value may seem very large, but it was desired to make the prediction extremely dependent on the most recent data, which is the most significant. The actual subroutine is included at the end of the paper but the simple mathematical model follows:

Let $A_{i,j,k}$ be the number of people in LOS cell k in quarter j of year i

For the four years of historical data calculate the loss rate for each quarter and each LOS independently

$$TR_{i,j,k} = \frac{A_{i,j,k} - A_{i+1,j,k+1}}{A_{i,j,k}}$$

The following series of data were smoothed for prediction

$\ldots, TR_{i,1,k}, TR_{i,2,k}, \ldots, TR_{i,4,k}, TR_{i+1,1,k}, \ldots\ldots$

For each LOS calculate the annual loss rate for prediction ($P_{i,k}$) using single exponential smoothing. The procedure is to iterate through four quarters of data which results in a single value for each year. The superscript (j), j = 1,2,3,4, is used to indicate the intermediate steps of

the procedure but need not be carried once the annual predicted loss rate has been calculated.

$$P_{i,k}^{(1)} = \alpha TR_{i,1,k} + (1.0 - \alpha)\ P_{i-1,k}$$

Where

$P_{i-1,k}$ is the final result from the previous year.

$$P_{i,k}^{(j)} = \alpha TR_{i,j,k} + (1.0 - \alpha)\ P_{i,k}^{(j-1)} \quad \text{for all } j = 2,3,4$$

$$P_{i,k} = P_{i,k}^{(4)}$$

Note that only one loss rate prediction is made for each year. This means that seasonal variations in the loss rate are not taken into account. Another approach would be to predict a loss rate separately for each quarter, the trade-off being that this procedure would require more years of data. This raises the question of whether using older data which takes into account seasonal variations would result in a better prediction than not using this older data but ignoring the seasonal variation. This is an area left for further study.

The prediction is now made for each quarter.

Let $T_{i,j,k}$ be the prediction for year i (first future year) quarter j and LOS k

$$T_{i,j,k} = A_{i-1,j,k-1}(1.0 - P_{i,k}) \qquad \text{for all } j = 1,2,3,4$$

As with the NAPPE model, this gives predictions for LOS cells 2-31. The calculation for cell 1 was done in a manner similar to the existing NAPPE model. The average proportion of the total force in cell 1 was calculated for the four years of data. For each quarter the following calculations were made:

Let C1AV be the number which would be required in cell 1 calculated by taking the above proportion of the total force requirement for the quarter being predicted.

$$\text{Let } C1P = Req - \sum_{k=2}^{31} T_{i,j,k} \qquad \text{total required minus the sum of cells 2-31}$$

Half of the difference between these two values was then allocated among cells 2-31 on the basis of the number projected for that cell.

For each cell calculate

$$ADJ_{i,j,k} = \frac{(C1AV - C1P)}{2} \, \frac{T_{i,j,k}}{\sum_{k=2}^{31} T_{i,j,k}}$$

$$T'_{i,j,k} = T_{i,j,k} + ADJ_{i,j,k}$$

The value for cell 1 is then the difference between the total required and the sum of the new predictions for cells 2-31.

2. <u>Results of SMOTHY</u>

The model, as described above, was then run to obtain one year predictions for the last ten years. The following table is a comparison of these results with outputs from NAPPE for the same time periods.

TABLE II

COMPARISON OF NAPPE AND SMOTHY PREDICTIONS
(costs are in millions of dollars)

Year	Actual Cost	NAPPE Prediction	% Error	SMOTHY Prediction	% Error
1965	1,362.1	1,360.0	.151	1,358.1	.209
1966	1,582.4	1,591.6	.501	1,582.0	.028
1967	1,720.7	1,736.5	.918	1,737.9	.997
1968	1,832.6	1,838.8	.337	1,836.9	.237
1969	2,009.2	2,016.7	.372	2,015.7	.322
1970	2,280.8	2,286.8	.262	2,280.9	.001
1971	2,264.8	2,258.2	.290	2,256.2	.379
1972	2,496.7	2,480.7	.643	2,483.6	.524
1973	2,683.6	2,678.1	.205	2,678.3	.196
1974	2,777.4	2,773.6	.138	2,772.1	.190
		Mean	.3817		.3083
		Mean 2	.3221		.2318

The value given in the table as mean is the mean of the absolute errors and the value called mean 2 is the same except the outlyer (1967) is left out of the calculation. Leaving this value out is not unreasonable when the events of 1967 are taken into consideration. It was during this year that the structure of the force saw tremendous change due to the Viet Nam buildup. Any model based on past data cannot predict the future when major policy decisions make that data inappropriate. This is a point where the analyst using the model must use reason when looking at its output, a point to be discussed later.

A close inspection of the preceeding table yields some surprising conclusions. Although the SMOTHY model does not predict uniformly better, it does significantly better for most years. This suggests that the complexities of the NAPPE model are not only unnecessary, but have a negative effect.

C. DOUBLE EXPONENTIAL SMOOTHING

The most important hypothesis tested was that single exponential smoothing is not the appropriate tool for modeling a time series which appears to have trends. As suggested by Brown [1], Goodman [2], and others, higher-order exponential smoothing is a valuable tool for modeling time series with underlying trends. Since the time series under consideration does not show any properties which would indicate anything

beyond a linear trend, only double smoothing was considered in SMOTH2.

1. <u>The SMOTH2 Subroutine</u>

As with the SMOTHY subroutine, the SMOTH2 subroutine makes use of only four years of historical data and only the TOTALNAV inventories. It also only smoothes the loss rates. These loss rates ($TR_{i,j,k}$) are calculated exactly the same as in SMOTHY, and the loss rate to be used for prediction ($D_{i,k}$) is calculated as described by Brown [1]. The single smoothed portion ($P_{i,k}$) is calculated exactly as before. The double smoothing term is calculated by simply smoothing the single smoothed value. The superscript notation is again used for the four iterations exactly as used to calculate $P_{i,k}$.

$$S^{(1)}_{i,k} = \alpha P^{(1)}_{i,k} + (1.0 - \alpha)\ S_{i-1,k}$$

$$S^{(j)}_{i,k} = \alpha P^{(j)}_{i,k} + (1.0 - \alpha)\ S^{(j-1)}_{i,k} \quad \text{for all } j = 2,3,4$$

$$S_{i,k} = S^{(4)}_{i,k}$$

These two values are then combined to pick up the linear trend and result in;

$$D_{i,k} = 2\ P_{i,k} - S_{i,k} + \frac{\alpha}{1.0 - \alpha}\ (P_{i,k} - S_{i,k})$$

Note that again only one value of the smoothed loss rate is calculated for each year and the same procedural question was left unanswered.

The prediction is then made for each quarter into the future exactly as in the SMOTHY subroutine

$$T_{i,j,k} = A_{i-1,j,k-1}(1.0 - D_{i,k})$$

As with the previously discussed models, this gives a prediction for LOS cells 2-31. The final adjustments used in SMOTH2 are exactly the same as used in SMOTHY.

2. Results of SMOTH2

As in the experiment with SMOTHY, it was obvious that the most recent data should be most heavily weighted. Therefore, an initial value of 0.4 was used for alpha. Since the impact of trend was the most important consideration of the research, other values of alpha were also tried. The following table is the result of these tests. In order to put the results in a form for analysis, the actual dollar values were not tabled but only the percentage errors. The elements of the table are the percentage error from the actual total cost of the force for NAPPE, NAPPE with the SMOTHY subroutine (these values are the same as Table II), and for NAPPE with the SMOTH2 subroutine using values of alpha of 0.2, 0.3, and 0.4.

TABLE III

COMPARISON OF ABSOLUTE PERCENTAGE ERRORS OF NAPPE, SMOTHY, AND SMOTH2

Year	NAPPE	SMOTHY	SMOTH2 $\alpha=.2$	SMOTH2 $\alpha=.3$	SMOTH2 $\alpha=.4$
1965	.151	.209	.242	.223	.215
1966	.501	.028	.015	.027	.018
1967	.918	.997	1.098	1.158	1.170
1968	.337	.237	.196	.098	.273
1969	.372	.322	.133	.033	.054
1970	.262	.001	.205	.076	.041
1971	.290	.379	.443	.423	.473
1972	.643	.524	.510	.527	.543
1973	.205	.196	.171	.164	.147
1974	.138	.190	.132	.068	.059
Mean	.3817	.3083	.3145	.2797	.2993
Mean 2	.3221	.2318	.2274	.1821	.2026
Mean 3	.2801	.1690	.1563	.0984	.1153

The values of mean and mean 2 have the same definition as in the preceeding table. The value mean 3 was calculated leaving out the values for 1967, 1971, and 1972. The reason for making this calculation will be discussed in detail later.

The first overview of the table would result in a conclusion that SMOTH2, with an alpha value of 0.3 seems to be a somewhat better model on the basis of the mean alone.

However, the mean is not the only significant feature. The most important observation is that SMOTH2 predicts extremely well for all years except 1967, 1971, and 1972. The reason for the poor prediction in 1967 has already been discussed. The reason for the poor prediction for 1971 and 1972 can be explained in the same manner except that the force structure was moving in the opposite direction. That is, these were the years of major policy changes resulting from the end of the Viet Nam War and the shrinking of the force.

Because of the linear trend, which is a part of SMOTH2, it must be expected to do poorly when the direction of the trend changes. This means that SMOTH2 should have more difficulty "turning the corner" when there are major policy changes. This does not mean that it is a poor model, but rather, some decision rule is required of the user when this occurs.

D. THE MATRIX GENERATION PROCEDURE

Although incomplete and inconclusive, a study of the Mosteller method, as used in this model, had some interesting results. It was found that, in general, the Mosteller procedure is extremely sensitive to the base matrix. It was found that changing the value in one cell of the base matrix resulted in changing the values in virtually every cell of the output matrix. There was no consistency found in these changes. The surprising result was that the changes in the output matrix were sometimes greater than the initial changes

in the base matrix. For example, changing one cell value in the base matrix by less than 5% could result in changes in the output matrix of greater than 5% in some cells.

This is extremely significant when considering its use in this model. The base matrix is calculated as a simple average of the last twelve quarterly inventories. This means that seasonal variations in the force structure are not taken into account and implies that a better base matrix may be possible.

A hypothesis was made that a better base matrix could be calculated by computing some very rough transition rates from one cell to another and these rates applied to the one year previous inventory. Experiments with this hypothesis showed some very promising results but were inconclusive. Continued study in this area may be of considerable value.

IV. CONCLUSIONS

The NAPPE model, in its present form of single exponential smoothing, does not appear to be the appropriate model. Single exponential smoothing has as an assumption that the time series is basically constant in time and the difference from the mean is caused by some random noise. This does not appear to be the case with LOS populations or with transition rates.

A recommended change to the model is to remove the complex SMOOTH and ADJSMO subroutines and replace them with the SMOTH2 subroutine, using an alpha of 0.3. It should be made clear to any intended user, however, that substantial changes in the enlisted force management would not be reflected in the prediction. The modelling approach should, in fact, be completely revised so that changes of this magnitude can be accounted for. Since pay grade totals are used to drive the force structure, the model is aware of impending changes in direction. This information is not currently being used in loss prediction by NAPPE.

In addition, the base matrix used in the Mosteller procedure could be estimated more carefully. Based on these preliminary experiments, this could result in a much better estimate of force structure, and hence a more accurate budget prediction. The determination of LOS cell 1 population remains somewhat ad hoc, as does the choice of a smoothing

constant of 0.3. While this study has demonstrated that a simpler approach to the prediction can be successfully taken, all other alternatives have not been thoroughly investigated.

APPENDIX A

This appendix is a rough documentation of the SMOOTH subroutine in the current NAPPE model.

Let

$A_{i,j,k}$ be the actual population for year i quarter j and LOS cell k

For each LOS k = 1, 2, ..., 30, calculate the transition rate for each year and quarter

$$TR_{i,j} = \frac{A_{i,j,k} - A_{i+1,j,k+1}}{A_{i,j,k}}$$

For each year i = 1, 2, ..., NYR-1 (NYR = last year of historical data) and each α = .05, .10, ..., .95, calculate:

$$P^1{}_5 = TR_{1,1}$$

Let n = 4i + j j = 1, 2, 3, 4

$$P^1{}_{n+1} = \alpha TR_{i,j} + (1.0 - \alpha)\, P^1{}_n$$

For each year then the predicted transition rate is

$$PRED_{i+2,\alpha} = P^1{}_{4i+1}$$

The relative error in this prediction is then the sum of the differences between the predicted and actual transition rates for the year

$$ER_{i+2,\alpha} = \sum_{j=1}^{4} \left| \frac{TR_{i+1,j} - P_{4i+1}}{1 - TR_{i+1,j}} \right|$$

Go to the INLINE subroutine to choose the best α.

INLINE Subroutine

For each year i = 1, 2, ..., NYR find the best α for predicting the following year.

For each α = .05, .10, ..., .95 calculate

$$EMIN_1 = ER_{i,\alpha}$$

$$EMIN_2 = \sum_{\ell=1}^{i} ER_{1,\alpha}$$

$$EMIN_3 = \sum_{\ell=1}^{i} (ER_{1,\alpha})^2$$

Select the α which gives the minimum value of $EMIN_1$, $EMIN_2$, $EMIN_3$ and call them α_1^*, α_2^*, α_3^*.

Calculate

$$SE_1 = \sum_{\ell=1}^{i} (ER_{1,\alpha_1^*})^2$$

$$SE_2 = \sum_{\ell=1}^{i} (ER_{1,\alpha_2^*})^2$$

$$SE_3 = \sum_{\ell=1}^{i} (ER_{1,\alpha_2^*})^2$$

Within the summation here, each α^* is the one which was selected for the given year

Select the minimum of these three values and the most recent α^* for that method is the α to be used to predict the following year. Call this value α_{i+1}

For each year i = 1, 2, ..., NYR and each quarter of the year, make the prediction based on the transition rate:

$$T_{i,j,k} = A_{i-1,j,k} (1.0 - PRED_{i,\alpha_i})$$

Now make a similar prediction based on population in each cell.

For each year i = 2, 3, ..., NYR and each α = .05, .10, ..., .95

$$P^1{}_8 = A_{1,4,k+1}$$

Let n = 4i+j   j = 1, 2, 3, 4

$$P^1{}_{n+1} = \alpha A_{i,j,k+1} + (1.0 - \alpha) P^1{}_n$$

$$PRED_{i+1,\alpha} = P^1_{4i+1}$$

The relative error in this prediction is then

$$ER_{i+1,\alpha} = \sum_{j=1}^{4} \left| \frac{A_{i+1,j,k+1} - P^1_{4i+1}}{A_{i+1,j,k+1}} \right|$$

Go to the INLINE subroutine to choose the best α. For each year i = 2, 3, ..., NYR and each quarter (the same value is predicted for all four quarters of a given year), make the prediction based on cell populations

$$P_{i,k} = PRED_{i,\alpha_i}$$

APPENDIX B

This appendix is a rough documentation of the ADJSMO subroutine in the current NAPPE model.

Define the five methods or techniques used in the subroutine:

1. $P(1)_{i,j} = T_{i,j,k}$ this is the predicted value for year i, quarter j calculated as the transition rate based prediction from SMOOTH using the TOTALNAV data.

2. $P(2)_{i,j} = T_{i,j,k}$ this is the sum of the predictions from SMOOTH made as above using USN and USNR data.

3. $P(3)_{i,j} = P_{i,k}$ this is the predicted value for year i, quarter j calculated as the population based prediction from SMOOTH using the TOTALNAV data.

4. $P(4)_{i,j} = P_{i,k}$ this is the sum of the predictions from SMOOTH made as above using USN and USNR data.

5. $P(5)_{i,j} = BWT \frac{P(1) + P(2)}{2} + (1.0 - BWT) \frac{P(3) + P(4)}{2}$

this is a weighted average of 1-4 where the value of BWT is the value which would have predicted best for the previous year.

For each LOS k = 2, 3, ..., 31

For each year i = 1, 2, ..., NYR calculate the cumulative square error for each method I = 1, 2, 3, 4, 5.

$$TEP(I)_i = \sum_{n=1}^{i} \sum_{j=1}^{4} \left(\frac{A_{i,j,k} - P(I)_{n,j}}{A_{i,j,k}}\right)^2$$

For each year calculate the cumulative square error for all values of BWT = .45, .50, ..., .95

$$ET2(BWT)_i = \sum_{n=1}^{i} \sum_{j=1}^{4} \left(\frac{A_{i,j,k} - X_{i,j}(BWT)}{A_{i,j,k}}\right)^2$$

where

$$X_{i,j}(BWT) = BWT \frac{P(1)_{i,j} + P(2)_{i,j}}{2} + (1.0 - BWT) \frac{P(3)_{i,j} + P(4)_{i,j}}{2}$$

Based on TEP select the best method, I, and based on ET2 select the best BWT which will then be used to calculate P(5) for the following year.

Calculate the cumulative absolute error for the entire period using the method which was selected as best for each year.

$$TER_k = \sum_{L=1}^{NYR} \sum_{S=1}^{4} \left| \frac{A_{i,j,k} - P(I_{n-1})_{i,j}}{A_{i,j,k}} \right|$$

where I_{n-1} was the best method for that year.

Make the initial prediction for the first future year

$$A_{NYR+1,j,k+1} + P(I_{NYR})_{NYR+1,j}$$

where I_{NYR} is the method selected best on the last year of data.

Let $IPG_{i,j}$ be the total force required

Calculate the average proportion of the total force in cell 1 for all years of data, call it PAV

For each quarter to be predicted, calculate

$$C1AV = (PAV)(IPG)$$

$$C1P = IPG - \sum_{k=2}^{31} A_{NYR+1,j,k}$$

These two values are the possible predictions for LOS 1. C1AV is based on the average proportion of the force in cell 1, while C1P is simply the difference between the total required force and the predictions for cells 2-31. Let

$$C1ADJ = \frac{C1AV + C1P}{2}$$

There is a test in the model to ensure that this average is between the values which would have been calculated using the largest and smallest proportions of the total population in cell 1 over the entire data base.

Take the difference between C1ADJ and C1P and allocate it among cells 2-31 according to the total error which was calculated for predicting that cell using the best method.

$$A^{*}_{i,j,k} = A_{i,j,k} + \frac{(C1P - C1ADJ)\,TER_{k-1} A_{i,j,k}}{\sum_{k=2}^{31} TER_{k-1}\, A_{i,j,k}}$$

Adjust cell 1 from these values

$$A^{*}_{i,j,1} = IPG_{i,j} - \sum_{k=1}^{31} A^{*}_{i,j,k}$$

The same basic procedure is used for predicting additional future years (up to 10) and the value of continuing the discussion is questionable.

```
// EXEC FORTCLG,REGION.GO=150K
//FORT.SYSPRINT DD DUMMY
//FORT.SYSIN DD *
      COMMCN NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $   APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &  TPOP2(30,4,30),
     $  IRED,ILM,ILY,
     $  IRT,IPUNCH,LHAVE(40)
      LOGICAL IRT,IPUNCH,LHAVE
      DIMENSION XMNLOS(10)
      INTEGER APOP
      DO 2 IYR=1,30
        DO 2 LOS=1,30
          PPOP2(IYR,LOS)=0.0
          DO 2 IQTR=1,4
            TPOP2(IYR,IQTR,LOS)=0.0
 2    CONTINUE
C
C  CALL SUBROUTINE TO READ & EDIT CONTROL CARDS
C
      IRED=0
      CALL RDCONT
      IF (IRT) WRITE (6,6000)
 6000 FORMAT (' CONTROL CARDS READ')
C
      IST=3
      CALL RDINV (ILM,ILY)
        IF (IRT) WRITE (6,6002) IST
 6002   FORMAT (' INVENTORIES READ ',I2)
      CALL SMOTH2
C  FOR EACH QTR FLIP-FLOP MATRIX & CSOT FORCE
C
      IMO=ILM
      IYR=ILY
      DO 20 NQ=1,NQTBP
      IMO=IMO+3
      IF (IMO.LT.13) GO TO 15
      IMO=3
      IYR=IYR+1
 15   CONTINUE
      CALL PNGPNG (IPG,APOP,NQ,NYR,NUMA,NUM)
      WRITE(7,7000) IMO,IYR,NQ,(J,(NUM(I,J),I=1,10),J=1,32)
 7000 FORMAT ('1',35X,'PREDICTION OF ALL NAVY ENLISTED INVENTORY FOR',I3
     $,'/',I2/45X,'PROJECTION LENGTH =',I3,' QUARTERS'//
     $3X,'LOS',12X,'E-1',8X,'E-2',8X,'E-3',8X,'E-4',8X,'E-5',8X,'E-6',
     $8X,'E-7',8X,'E-8',8X,'E-9',6X,'TOTAL'/32(/4X,I2,4X,10I11))
       IF (IPUNCH) WRITE(4,5000) (ILM,ILY,IMO,IYR,(NUM(I,J),I=1,10),J,
```

```
     $   J=1,32)
 5000 FORMAT (4I2,10I7,I2)
      WRITE (7,7001)
      DO 30  MN=1,10
      X=0.0
         DO 31  LOS=1,31
   31    X = X + (LOS-.5)*NUM(MN,LOS)
      M = NUM(MN,32)
      IF (M.EQ.0)  XMNLOS(MN)=0.0
   30 IF (M.NE.0)  XMNLOS(MN) = X/M
      WRITE (7,4000) (XMNLOS(MN),MN=1,10)
 4000 FORMAT('0',8X,'E-1',9X,'E-2',9X,'E-3',9X,'E-4',9X,'E-5',9X,'E-6',
     $9X,'E-7',9X,'E-8',9X,'E-9',7X,'TOTAL'//10F12.3)
 7001 FORMAT('0',20X,'MEAN LOS DISTRIBUTION FOR THE QUARTER')
      CALL COSTPG (IMO,IYR)
 20   CONTINUE
      RETURN
      END
C
C
      SUBROUTINE RDCONT
C
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),
     $ NUMG(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $    APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &  TPOP2(30,4,30),
     $  IRED,ILM,ILY,
     $  IRT,IPUNCH,LHAVE(40)
      LOGICAL IRT,IPUNCH,LHAVE
      INTEGER APOP
C
C
      DIMENSION HEAD(19)
C
C  WRITE HEADER
C
      DO 10 I=1,21
        READ  (11,1100) (HEAD(J),J=1,19)
 10     WRITE (7,7000)(HEAD(J)   ,J=1,19)
 1100 FORMAT(19A4)
 7000 FORMAT (25X,19A4)
C
C  READ CONTROL INFO
C
      READ (5,5000) ILM,ILY,NQTBP,IRTT,IPNCH
 5000 FORMAT (3I2,72X,2I1)
C
      IRT=.FALSE.
```

```
      IF (IRTT.EQ.1) IRT=.TRUE.
      IPUNCH=.FALSE.
      IF (IPNCH.EQ.1) IPUNCH=.TRUE.
      IF (NQTBP.LE.40) GO TO 5
      WRITE (6,6000) NQTBP
 6000 FORMAT (2X,I2,' QUARTERS REQUESTED; ONLY 40 CAN BE PROJECTED')
      NQTBP = 40
C
 5    WRITE (7,7002) NQTBP,ILM,ILY,IRTT,IPNCH
C
 7002 FORMAT (6(/),20X,'NOTES:'//26X,'THE',I3,' QUARTERLY PROJECTIONS MA
     $DE ARE BASED ON HISTORICAL DATA THRU',I3,'/19',I2,18X,2I1,/25X,
     $'USER INPUT PAY GRADE TOTALS USED IN PROJECTIONS ARE:'/20X,'FOR',
     $5X,'E-1',5X,'E-2',5X,'E-3',5X,'E-4',5X,'E-5',5X,'E-6',5X,'E-7',5X,
     $'E-8',5X,'E-9',3X,'TOTAL')
C
      DO 30 I=1,NQTBP
 30     LHAVE(I)=.FALSE.
C
      NLKQ=ILM/3+ILY*4
C
      NYR=(NLKQ-228)/4
      NCTS=(NLKQ-228-4*NYR)*32
 80   IF (NYR.LE.20) GO TO 85
      NYR=NYR-1
      NCTS=NCTS+128
      GO TO 80
 85   NYRM1=NYR-1
      NYRP1=NYR+1
C
C
 40   READ (5,5001,END=60) IMO,IYR,ILJ
 5001 FORMAT (2X,2I2,2X,10I7)
      NQPG=IMO/3+IYR*4-NLKQ
      IF ((NQPG.LT.1).OR.(NQPG.GT.40)) GO TO 45
      IF (LHAVE(NQPG)) GO TO 45
      DO 50 I=1,10
 50     IPG(NQPG,I)=ILJ(I)
      WRITE (7,7004) IMO,IYR,ILJ
 7004 FORMAT (19X,I2,'/',I2,10I8)
      LHAVE(NQPG)=.TRUE.
      GO TO 40
C
 45   WRITE (6,6001) IMO,IYR,ILJ
 6001 FORMAT (' PAY GRADE CARD:',2I3,10I8,2X,'NOT USED')
      GO TO 40
C
C CHECK FOR SUFFICIENT PAY GRADE TOTALS
```

```
C
 60   DO 65 I=1,NQTBP
         IF (.NOT.LHAVE(I)) GO TO 70
 65   CONTINUE
      RETURN
C
 70   IMO=ILM-2
      IF (ILM.GT.6) IMO=(18-ILM)/3
 75   IF (.NOT.LHAVE(IMO)) GO TO 90
      IF (IMO.GE.NQTBP) RETURN
      IMO=IMO+4
      GO TO 75
C
 90   WRITE (6,6002)
 6002 FORMAT (' INSUFFICIENT PAY GRADE TOTALS INPUT')
      RETURN
      END
C
C
      SUBROUTINE RDINV
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $    APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &   TPOP2(30,4,30),
     $   IRED,ILM,ILY,
     $   IRT,IPUNCH,LHAVE(40)
      LOGICAL IRT,IPUNCH,LHAVE
      INTEGER APOP
      LOGICAL FJUNE
      IF (IST.NE.3) GO TO 4
      FJUNE=.FALSE.
      NYRM2=NYR-2
      DO 2 I=1,10
         DO 2 J=1,32
 2          NUMA(I,J)=0
 4    IF (NCTS.EQ.0) GO TO 6
      DO 5 I=1,NCTS
 5       READ (IST,1000)
 1000 FORMAT (80X)
 6    CONTINUE
      DO 50 IYR=1,NYR
         DO 50 IQTR=1,4
            READ (IST,1005,END=100) IMO,JYR,((NUM(I,J),I=1,10),J=1,32)
 1005       FORMAT (2X,2I2,2X,10I7,31(/8X,10I7))
            DO 10 J=1,32
 10            APOP(IYR,IQTR,J)=NUM(10,J)
            IF (IST.NE.3) GO TO 50
            IF (IYR.LT.NYRM2) GO TO 50
```

```
      DO 20 I=1,10
        DO 20 J=1,32
20        NUMA(I,J)=NUMA(I,J)+NUM(I,J)
      IF (IYR.NE.NYR) GO TO 50
      IF (FJUNE) CALL COSTPG(IMO,JYR)
      IF (IMO.NE.6) GO TO 50
      FJUNE=.TRUE.
      DO 40 I=1,10
        ILJ(I)=NUM(I,32)
        DO 40 J=1,32
40        NUMO(I,J)=NUM(I,J)
50    CONTINUE
      REWIND IST
      RETURN
100   WRITE (6,6000) IST
6000  FORMAT (' INSUFFICIENT DATA ON FILE',I2)
      RETURN
      END
C
C
C
C
      SUBROUTINE SMOTH2
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $    APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &   TPOP2(30,4,30),
     $   IRED,ILM,ILY,
     $   IRT,IPUNCH,LHAVE(40)
      DIMENSION TR(20,4),PRED(20),TPOP(32,4,32)
      INTEGER APOP
      NQ=0
      LNY=NYR
      INY=NYR-4
      JNY=NYRM1
      KNY=NYRP1
 202  CONTINUE
      DO 200 LOS=1,30
      LOSP1=LOS+1
         DO 10 IYR=INY,JNY
              DO 10 IQTR=1,4
              SPOP=APUP(IYR,IQTR,LOS)
              IF(LOS.EQ.30) SPOP=SPOP+APOP(IYR,IQTR,31)
   10 TR(IYR,IQTR)=(SPOP-APOP(IYR+1,IQTR,LOSP1))/SPOP
      ALF=.4
      S1=TR(INY,1)
      S2=S1
      ISQ=2
```

```
      DO 30 IYR=INY,JNY
         DO 20 IQTR=ISQ,4
         S1=ALF*TR(IYR,IQTR) + (1.0-ALF)*S1
         S2=ALF*S1 + (1.0-ALF)*S2
         ISQ=1
   20    CONTINUE
   30 CONTINUE
      AHAT=2.0*S1 - S2
      BHAT=(ALF/(1.0-ALF))*(S1-S2)
      PRED(IYR+2)=AHAT+BHAT
      DO 75 IQTR=1,4
   75 TPOP(KNY,IQTR,LOSP1)=APOP(LNY,IQTR,LOS)*(1.0-PRED(KNY))
  200 CONTINUE
      PAV=0.0
      DO 120 IYR=INY,LNY
         DO 120 IQ=1,4
         R=APOP(IYR,IQ,1)
         R=R/APOP(IYR,IQ,32)
  120 PAV=PAV+R
      PAV=PAV/20.
      IYR=KNY
      DO 160 IQ=1,4
      NQ=NQ+1
      TP=0.0
      C1AV=PAV*IPG(NQ,10)
      C1P=IPG(NQ,10)
         DO 130 LOS=2,31
         C1P=C1P-TPCP(IYR,IQ,LOS)
  130 TP=TP+TPOP(IYR,IQ,LOS)
         C1=.5*(C1P+C1AV)
      APOP(IYR,IQ,1)=IPG(NQ,10)
         DO 135 LOS=2,31
         ADJ=(C1P-C1)*TPOP(IYR,IQ,LOS)/TP
         APOP(IYR,IQ,LOS)=TPOP(IYR,IQ,LOS) + ADJ
  135 APOP(IYR,IQ,1)=APOP(IYR,IQ,1) - APOP(IYR,IQ,LOS)
  160 CONTINUE
      IF(NQ.EQ.NQTBP) GO TO 201
      LNY=LNY+1
      INY=INY+1
      JNY=JNY+1
      KNY=KNY+1
      GO TO 202
  201 RETURN
      END
C
C
      SUBROUTINE PNGPNG (IPG,APOP,NQ,NYR,NUM,NUM2)                      00005020
C   SUBROUTINE PNGPNG FINDS INVENTORY NUM2 THAT HAS SIDE TOTALS ISIDE   00005030
```

```
C AND FOOT TOTALS IFOOT THAT HAS APPROXIMATELY THE SAME DISTRIBUTION
C AS NUM.
      DIMENSION ISIDE(32),IFOOT(10),NUM(10,32),NUM2(10,32),XPRE(32,10)
      DIMENSION ID(31),JD(9)
      INTEGER APOP
      DIMENSION IPG(40,10),APOP(32,4,32)
      DO 1 I=1,10
 1    IFOOT(I)=IPG(NQ,I)
      J=(NQ+3)/4
      K=NQ-J*4+4
      J=J+NYR
      DO 2 I=1,31
 2    ISIDE(I)=APOP(J,K,I)
      DO 10 J=1,9
      GOT=IFOOT(J)
        WANT=NUM(J,32)
      IF (WANT.LE.0.) GO TO 5
        WANT=GOT/WANT
 5    CONTINUE
        DO 10 I=1,31
          GOT=NUM(J,I)
 10       XPRE(I,J)=GOT*WANT
      DO 90 IC=1,200
 20   SR=0.
      TE=0.
      DO 50 I=1,31
        GOT=0.
        DO 30 J=1,9
            X=XPRE(I,J)-NUM(J,I)
      TE=TE+ABS(X)
 30       GOT=GOT+XPRE(I,J)
      WANT=ISIDE(I)
        R=ABS(WANT-GOT)
        IF (R.LE.1.) GO TO 50
        SR=SR+R
      IF (GOT.LE.0.) GO TO 50
        R=WANT/GOT
        DO 40 J=1,9
 40       XPRE(I,J)=XPRE(I,J)*R
 50   CONTINUE
      IF (SR.LE.1.) GO TO 100
      SR=0.
      TE=0.
      DO 80 J=1,9
        GOT=0.
        DO 60 I=1,31
            X=XPRE(I,J)-NUM(J,I)
      TE=TE+ABS(X)
```

```
60          GOT=GOT+XPRE(I,J)
      WANT=IFOOT(J)
        R=ABS(WANT-GOT)
        IF (R.LE.1.) GO TO 80
        SR=SR+R
      IF (GOT.LE.0.) GO TO 80
        R=WANT/GOT
        DO 70 I=1,31
70        XPRE(I,J)=XPRE(I,J)*R
80    CONTINUE
1005  FORMAT (8X,I4,2F15.1)
      IF (SR.LE.1.) GO TO 100
90    CONTINUE
      WRITE (6,1000)
1000  FORMAT (' **** ERROR : PING PONG DOES NOT CONVERGE.')
100   DO 110 I=1,31
          NUM2(10,I)=0
        DO 110 J=1,9
          NUM2(J,I)=XPRE(I,J)+.5
110       NUM2(10,I)=NUM2(10,I)+NUM2(J,I)
      DO 120 J=1,10
        NUM2(J,32)=0
        DO 120 I=1,31
120       NUM2(J,32)=NUM2(J,32)+NUM2(J,I)
      RETURN
      END
C
C
C
      SUBROUTINE COSTPG (IMO,IYR)
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $   APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &  TPOP2(30,4,30),
     $  IRED,ILM,ILY,
     $  IRT,IPUNCH,LHAVE(40)
      LOGICAL IRT,IPUNCH,LHAVE
      INTEGER APOP
      LOGICAL LNR
      DIMENSION RATE2(10)
      LNR=.TRUE.
      J=IMO/3
      DO 5 I=1,10
        COST(I,J)=0.0
        RATE2(I)=0.0
5     CONTINUE
      IQED=IYR*360+IMO*30+30
10    IF (IQED.LE.IRED) GO TO 20
```

```
      P=IQED-IRED
      P=P/90.0
      READ (8,8000,END=500) IRM,IRD,IRY
      IRED=IRY*360+IRM*30+IRD
      IF (P.LT.1) GO TO 25
      READ (8,8005,END=500) ((RATE(I,K),I=1,10),K=1,31)
      GO TO 10
20    LNR=.FALSE.
      P=0.0
25    PM1=1.0-P
      DO 30 I=1,31
        IF (LNR) READ (8,8005,END=500) RATE2
        DO 30 K=1,9
          COST(K,J)=COST(K,J)+(NUMO(K,I)+NUM(K,I))*(RATE(K,I)*PM1+
     $        RATE2(K)*P)*1.5
          IF (LNR) RATE(K,I)=RATE2(K)
30    CONTINUE
      A=0.0
      DO 40 I=1,9
        A=A+COST(I,J)
        DO 40 K=1,31
          NUMO(I,K)=NUM(I,K)
40    CONTINUE
      COST(10,J)=A
      WRITE (7,4005) IMO,IYR
      WRITE (7,4000) (COST(I,J),I=1,10)
      IF (IPUNCH) WRITE (4,7000) (COST(I,J),I=1,10)
      IF (IMO.NE.6) RETURN
      DO 50 J=1,10
        A=0.0
        DO 45 I=1,4
45        A=A+COST(J,I)
50      RATE2(J)=A
      WRITE(7,4010) IYR
7000  FORMAT (5F12.0)
      IF (IPUNCH) WRITE (4,7000) RATE2
      WRITE (7,4000) RATE2
4000  FORMAT ('0',8X,'E-1',9X,'E-2',9X,'E-3',9X,'E-4',9X,'E-5',9X,'E-6',
     $ 9X,'E-7',9X,'E-8',9X,'E-9',7X,'TOTAL'//10F12.0)
4005  FORMAT ('0',20X,'PROJECTED ENLISTED BASIC PAY FOR QUARTER ENDING',
     $ I3,'/',I2)
4010  FORMAT ('0',20X,'PROJECTED ENLISTED BASIC PAY FOR FISCAL YEAR 19',
     $ I2)
      RETURN
8000  FORMAT (3I2)
8005  FORMAT (8X,10F7.2)
500   WRITE (6,6000)
6000  FORMAT ('  $$ ERROR')
```

```
      RETURN                                                            00006470
      END                                                               00006480
      SUBROUTINE PREDPG                                                 00006490
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),         00006500
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),       00006510
     $   APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),        00006520
     &  TPOP2(30,4,30),                                                 00006530
     $  IRED,ILM,ILY,                                                   00006540
     $  IRT,IPUNCH,LHAVE(40)                                            00006550
      LOGICAL IRT,IPUNCH,LHAVE                                          00006560
      INTEGER APOP                                                      00006570
      DIMENSION ST(3,9),PRE(3,10)                                       00006580
      DATA ST/1.1248,.8539,1.0040,1.0146,1.0555,1.0426,1.0454,.9768,    00006590
     11.0376,.9380,.9848,.9557,.9361,.9712,.9389,.9619,.9836,           00006600
     2.9687,1.0086,.9355,.9831,.9661,1.0284,1.0191,.9705,1.0142,1.0054/ 00006610
      NJ=ILM/3                                                          00006620
      IF (NJ.NE.1) NJ=6-NJ                                              00006630
5     DO 10 I=1,NQTBP                                                   00006640
        IF (.NOT.LHAVE(I)) GO TO 20                                     00006650
10    CONTINUE                                                          00006660
      RETURN                                                            00006670
20    CONTINUE                                                          00006680
      IF (.NOT.LHAVE(NJ)) GO TO 100                                     00006690
      IF (I.LT.NJ) GO TO 40                                             00006700
      GO TO 66                                                          00006710
40    PRE(1,10)=0.                                                      00006720
      PRE(2,10)=0.                                                      00006730
      PRE(3,10)=0.                                                      00006740
      DO 60 K=1,9                                                       00006750
        PRE(2,K)=(ILJ(K)+IPG(NJ,K))/2                                   00006760
      PRE(1,K)=(ILJ(K)+PRE(2,K))/2                                      00006770
        PRE(3,K)=(PRE(2,K)+IPG(NJ,K))/2                                 00006780
        DO 50 J=1,3                                                     00006790
          PRE(J,K)=PRE(J,K)*ST(J,K)                                     00006800
50        PRE(J,10)=PRE(J,10)+PRE(J,K)                                  00006810
60      CONTINUE                                                        00006820
      DO 65 K=1,3                                                       00006830
        I=NJ-K                                                          00006840
        IF (I.LE.0) GO TO 65                                            00006850
        IF (LHAVE(I)) GO TO 65                                          00006860
        LHAVE(I)=.TRUE.                                                 00006870
        DO 62 J=1,10                                                    00006880
62        IPG(I,J)=PRE(4-K,J)+.5                                        00006890
65    CONTINUE                                                          00006900
66    CONTINUE                                                          00006910
      DO 70 I=1,10                                                      00006920
70      ILJ(I)=IPG(NJ,I)                                                00006930
      NJ=NJ+4                                                           00006940
```

```
      GO TO 5
100   WRITE (6,1000)
1000  FORMAT ('  $$$$ERROR : JUNE PAY GRADE TOTALS MUST BE INPUT')
      RETURN
      END
```

```
      CALL SMOTHY
      SUBROUTINE SMOTHY
      COMMON NYR,NYRM1,NYRP1,IST,NQTBP,NCTS,ILJ(10),NUM(10,32),          00
     $ NUMO(10,32),NUMA(10,32),IPG(40,10),RATE(10,32),COST(10,4),
     $   APOP(32,4,32),PPOP1(30,30),PPOP2(30,30),TPOP1(30,4,30),
     &  TPOP2(30,4,30),
     $  IRED,ILM,ILY,
     $  IRT,IPUNCH,LHAVE(40)
      DIMENSION TR(20,4),PRED(20),TPOP(32,4,32)
      INTEGER APOP
      NQ=0
      LNY=NYR
      INY=NYR-4
      JNY=NYRM1
      KNY=NYRP1
  202 CONTINUE
      DO 200 LOS=1,30
      LOSP1=LOS+1
         DO 10 IYR=INY,JNY
              DO 10 IQTR=1,4
              SPOP=APOP(IYR,IQTR,LOS)
              IF(LOS.EQ.30) SPOP=SPOP+APOP(IYR,IQTR,31)
   10 TR(IYR,IQTR)=(SPOP-APOP(IYR+1,IQTR,LOSP1))/SPOP
      ALF=.4
      P1=TR(INY,1)
      ISQ=2
      DO 30 IYR=INY,JNY
         DO 20 IQTR=ISQ,4
         P1=ALF*TR(IYR,IQTR)+(1.0-ALF)*P1
      ISQ=1
   20 CONTINUE
   30 CONTINUE
      PRED(IYR+2)=P1
      DO 75 IQTR=1,4
   75 TPOP(KNY,IQTR,LOSP1)=APOP(LNY,IQTR,LOS)*(1.0-PRED(KNY))
  200 CONTINUE
      PAV=0.0
      DO 120 IYR=INY,LNY
         DO 120 IQ=1,4
         R=APOP(IYR,IQ,1)
         R=R/APOP(IYR,IQ,32)
  120 PAV=PAV+R
      PAV=PAV/20.
      IYR=KNY
      DO 160 IQ=1,4
      NQ=NQ+1
      TP=0.0
      C1AV=PAV*IPG(NQ,10)
```

```
      C1P=IPG(NQ,10)
         DO 130 LOS=2,31
         C1P=C1P-TPOP(IYR,IQ,LOS)
  130 TP=TP+TPOP(IYR,IQ,LOS)
         C1=.5*(C1P+C1AV)
      APOP(IYR,IQ,1)=IPG(NQ,10)
         DO 135 LOS=2,31
         ADJ=(C1P-C1)*TPOP(IYR,IQ,LOS)/TP
         APOP(IYR,IQ,LOS)=TPOP(IYR,IQ,LOS) + ADJ
  135 APOP(IYR,IQ,1)=APOP(IYR,IQ,1) - APOP(IYR,IQ,LOS)
  160 CONTINUE
      IF(NQ.EQ.NQTBP) GO TO 201
      LNY=LNY+1
      INY=INY+1
      JNY=JNY+1
      KNY=KNY+1
      GO TO 202
  201 RETURN
      END
```

LIST OF REFERENCES

1. Brown, R.G., Smoothing, Forecasting and Prediction, p. 97-144, Prentice-Hall, Inc., 1963.

2. Goodman, M.L., "A New Look at Higher-Order Exponential Smoothing for Forecasting", Operations Research, p. 880-888, July-August 1974.

3. Mosteller, F., "Association and Estimation in Contingency Tables", Journal of the American Statistical Association, p. 1-4, March 1968.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	2
4.	Chief of Naval Personnel Attention: Mr. R.K. Lehto PERS 2x Department of the Navy Washington, D.C. 20370	2
5.	Mr. Joe Silverman Manpower Systems Naval Personnel Research and Development Center San Diego, California 92152	5
6.	Assoc. Professor R.W. Butterworth, Code 55Bd Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	10
7.	Professor R.R. Read, Code 55Re Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
8.	Assoc. Professor P.R. Milch, Code 55Mh Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
9.	Professor K.T. Marshall, Code 55Mt Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1

10. Chief of Naval Personnel 1
Attention: LCDR J.H. SMITH, PERS 21221
Department of the Navy
Washington, D.C. 20370

11. LT Allan R. Walker 1
VAW 113
FPO San Francisco, California 96601

Thesis
W22037 Walker
c.1
An analysis of the Naval Personnel Pay Predictor (Enlisted Model).